



SEQUENCE LISTING

<110> Kincaid, Randall L.
Angov, Evalina
Lyon, Jeffrey A.
Veritas, Inc.

<120> Improved Protein Expression by Codon Harmonization and
Translational Attenuation

<130> 016873-000400US

<140> US 10/677,641
<141> 2003-10-01

<150> US 60/369,741
<151> 2002-04-01

<150> US 60/379,688
<151> 2002-05-09

<150> US 60/425,719
<151> 2002-11-12

<150> WO PCT/US03/10384
<151> 2003-04-01

<160> 10

<170> PatentIn Ver. 2.1

<210> 1
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR
amplification primer FVO-PCR1

<400> 1
gggtcggtac catggcagta actccttccg taattgat

38

<210> 2
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR
amplification primer FVO-PCR2

<400> 2
ggatcagatg cggccgctta actgcagaaa ataccatcga aaagtggaa

48

<210> 3
<211> 23
<212> DNA
<213> Artificial Sequence

```

<220>
<223> Description of Artificial Sequence:PCR
      primer-directed mutagenesis overlapping
      oligonucleotide EA3

<400> 3
taaaaaatat ataaacgaca aac                                23

<210> 4
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR
      primer-directed mutagenesis overlapping
      oligonucleotide EA5

<400> 4
aaaaggaaag atatttctca ttt                                23

<210> 5
<211> 71
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:synthesized
      sense strand oligonucleotide EA485-CDFVO

<400> 5
cgcagttact ccatctgtta ttgataatat tctttctaaa attgaaaacg aatatgagg 60
tttatattta a                                71

<210> 6
<211> 79
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:synthesized
      sense strand oligonucleotide EA493-CDFVO

<400> 6
ggttttaaat ataaaacctc atattcgtt tcaatttttag aaagaatatt atcaataaca 60
gatggagtaa ctgcggtag                                79

<210> 7
<211> 1065
<212> DNA
<213> Plasmodium falciparum

<220>
<223> native codon MSP1-42

```

<220>
 <221> CDS
 <222> (1)..(1065)
 <223> MSP1-42

<400> 7
 gca gta act cct tcc gta att gat aac ata ctt tct aaa att gaa aat 48
 Ala Val Thr Pro Ser Val Ile Asp Asn Ile Leu Ser Lys Ile Glu Asn
 1 5 10 15

gaa tat gag gtt tta tat tta aaa cct tta gca ggt gtt tat aga agt 96
 Glu Tyr Glu Val Leu Tyr Leu Lys Pro Leu Ala Gly Val Tyr Arg Ser
 20 25 30

tta aaa aaa caa tta gaa aat aac gtt atg aca ttt aat gtt aat gtt 144
 Leu Lys Lys Gln Leu Glu Asn Asn Val Met Thr Phe Asn Val Asn Val
 35 40 45

aag gat att tta aat tca cga ttt aat aaa cgt gaa aat ttc aaa aat 192
 Lys Asp Ile Leu Asn Ser Arg Phe Asn Lys Arg Glu Asn Phe Lys Asn
 50 55 60

gtt tta gaa tca gat tta att cca tat aaa gat tta aca tca agt aat 240
 Val Leu Glu Ser Asp Leu Ile Pro Tyr Lys Asp Leu Thr Ser Ser Asn
 65 70 75 80

tat gtt gtc aaa gat cca tat aaa ttt ctt aat aaa gaa aaa aga gat 288
 Tyr Val Val Lys Asp Pro Tyr Lys Phe Leu Asn Lys Glu Lys Arg Asp
 85 90 95

aaa ttc tta agc agt tat aat tat att aag gat tca ata gat acg gat 336
 Lys Phe Leu Ser Ser Tyr Asn Tyr Ile Lys Asp Ser Ile Asp Thr Asp
 100 105 110

ata aat ttt gca aat gat gtt ctt gga tat tat aaa ata tta tcc gag 384
 Ile Asn Phe Ala Asn Asp Val Leu Gly Tyr Tyr Lys Ile Leu Ser Glu
 115 120 125

aaa tat aaa tca gat tta gat tca att aaa aaa tat ata aac gac aaa 432
 Lys Tyr Lys Ser Asp Leu Asp Ser Ile Lys Lys Tyr Ile Asn Asp Lys
 130 135 140

caa ggt gaa aat gag aaa tac ctt ccc ttt tta aac aat att gag acc 480
 Gln Gly Glu Asn Glu Lys Tyr Leu Pro Phe Leu Asn Asn Ile Glu Thr
 145 150 155 160

tta tat aaa aca gtt aat gat aaa att gat tta ttt gta att cat tta 528
 Leu Tyr Lys Thr Val Asn Asp Lys Ile Asp Leu Phe Val Ile His Leu
 165 170 175

gaa gca aaa gtt cta aat tat aca tat gag aaa tca aac gta gaa gtt 576
 Glu Ala Lys Val Leu Asn Tyr Thr Tyr Glu Lys Ser Asn Val Glu Val
 180 185 190

aaa ata aaa gaa ctt aat tac tta aaa aca att caa gac aaa ttg gca 624
 Lys Ile Lys Glu Leu Asn Tyr Leu Lys Thr Ile Gln Asp Lys Leu Ala
 195 200 205

gat ttt aaa aaa aat aac aat ttc gtt gga att gct gat tta tca aca 672
 Asp Phe Lys Lys Asn Asn Phe Val Gly Ile Ala Asp Leu Ser Thr
 210 215 220

gat tat aac cat aat aac tta ttg aca aag ttc ctt agt aca ggt atg	720
Asp Tyr Asn His Asn Asn Leu Leu Thr Lys Phe Leu Ser Thr Gly Met	
225 230 235 240	
gtt ttt gaa aat ctc gct aaa acc gtt tta tct aat tta ctt gat gga	768
Val Phe Glu Asn Leu Ala Lys Thr Val Leu Ser Asn Leu Leu Asp Gly	
245 250 255	
aac ttg caa ggt atg tta aac att tca caa cac caa tgc gta aaa aaa	816
Asn Leu Gln Gly Met Leu Asn Ile Ser Gln His Gln Cys Val Lys Lys	
260 265 270	
caa tgt cca caa aat tct gga tgt ttc aga cat tta gat gaa aga gaa	864
Gln Cys Pro Gln Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu	
275 280 285	
gaa tgt aaa tgt tta tta aat tac aaa caa gaa ggt gat aaa tgt gtt	912
Glu Cys Lys Cys Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val	
290 295 300	
gaa aat cca aat cct act tgt aac gaa aat aat ggt gga tgt gat gca	960
Glu Asn Pro Asn Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala	
305 310 315 320	
gat gcc aaa tgt acc gaa gaa gat tca ggt agc aac gga aag aaa atc	1008
Asp Ala Lys Cys Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile	
325 330 335	
aca tgt gaa tgt act aaa cct gat tct tat cca ctt ttc gat ggt att	1056
Thr Cys Glu Cys Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile	
340 345 350	
ttc tgc agt	1065
Phe Cys Ser	
355	

<210> 8
 <211> 355

<212> PRT
 <213> Plasmodium falciparum

<220>
 <223> native codon MSP1-42

<400> 8
 Ala Val Thr Pro Ser Val Ile Asp Asn Ile Leu Ser Lys Ile Glu Asn
 1 5 10 15

Glu Tyr Glu Val Leu Tyr Leu Lys Pro Leu Ala Gly Val Tyr Arg Ser
 20 25 30

Leu Lys Lys Gln Leu Glu Asn Asn Val Met Thr Phe Asn Val Asn Val
 35 40 45

Lys Asp Ile Leu Asn Ser Arg Phe Asn Lys Arg Glu Asn Phe Lys Asn
 50 55 60

Val Leu Glu Ser Asp Leu Ile Pro Tyr Lys Asp Leu Thr Ser Ser Asn
 65 70 75 80

Tyr Val Val Lys Asp Pro Tyr Lys Phe Leu Asn Lys Glu Lys Arg Asp
 85 90 95
 Lys Phe Leu Ser Ser Tyr Asn Tyr Ile Lys Asp Ser Ile Asp Thr Asp
 100 105 110
 Ile Asn Phe Ala Asn Asp Val Leu Gly Tyr Tyr Lys Ile Leu Ser Glu
 115 120 125
 Lys Tyr Lys Ser Asp Leu Asp Ser Ile Lys Lys Tyr Ile Asn Asp Lys
 130 135 140
 Gln Gly Glu Asn Glu Lys Tyr Leu Pro Phe Leu Asn Asn Ile Glu Thr
 145 150 155 160
 Leu Tyr Lys Thr Val Asn Asp Lys Ile Asp Leu Phe Val Ile His Leu
 165 170 175
 Glu Ala Lys Val Leu Asn Tyr Thr Tyr Glu Lys Ser Asn Val Glu Val
 180 185 190
 Lys Ile Lys Glu Leu Asn Tyr Leu Lys Thr Ile Gln Asp Lys Leu Ala
 195 200 205
 Asp Phe Lys Lys Asn Asn Phe Val Gly Ile Ala Asp Leu Ser Thr
 210 215 220
 Asp Tyr Asn His Asn Asn Leu Leu Thr Lys Phe Leu Ser Thr Gly Met
 225 230 235 240
 Val Phe Glu Asn Leu Ala Lys Thr Val Leu Ser Asn Leu Asp Gly
 245 250 255
 Asn Leu Gln Gly Met Leu Asn Ile Ser Gln His Gln Cys Val Lys Lys
 260 265 270
 Gln Cys Pro Gln Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu
 275 280 285
 Glu Cys Lys Cys Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val
 290 295 300
 Glu Asn Pro Asn Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala
 305 310 315 320
 Asp Ala Lys Cys Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Ile
 325 330 335
 Thr Cys Glu Cys Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile
 340 345 350
 Phe Cys Ser
 355

<210> 9
 <211> 1065
 <212> DNA
 <213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Plasmodium
falciparum MSP1-42 with harmonized codons for
expression in E. coli cl. 2

<400> 9
gctgttacgc catcggttat tgacaatatac cttagcaaaaa ttgaaaacga atacgagggtt 60
ctgtacctga aaccactggc tggcgtttac cgttctctga aaaaacagct ggaaaacaat 120
gttatgacct tcaacgttaa cgttaaggac attctgaaca gccgcttcaa caaacgcgaa 180
aacttcaaaa acgttctgga aagcgacctg attccataca aagacctgac cagctctaac 240
tacgttgtca aagaccata caaattcctc aacaaagaaa aacgtgacaa attcctgtcg 300
tcttacaact acattaagga cagcatcgac acggacatca acttcgctaa cgacgttctc 360
ggctactaca aaatcctgtc ggagaaatac aaaagcgacc tggacagcat taaaaaatac 420
atcaatgata aacagggcga aaacgagaaa tatctccctt tcctgaataa cattgagacg 480
ctgtacaaaaa ccgttaacga caaaattgac ctgttcgtta ttcacctgga agctaaagtt 540
ctcaactaca cctacgagaa aagcaatgtt gaagttaaaa tcaaagaact caactatctg 600
aaaaccattc aggataaaact cgctgacttc aaaaaaaaca ataacttcgt tggcattgt 660
gacctgagca ccgactacaa tcacaacaat ctgctcacca agttcccttc taccggcatg 720
gttttcgaaa acctcgctaa aacggttctg agcaacctgc tcgacggcaa tctccagggc 780
atgctgaata ttagccagca tcagtgtgtt aaaaaacagt gcccacagaa cagcggctgc 840
ttccgtcacc tggacgaacg tgaagaatgc aaatgcctgc tgaactataa acaggaaggc 900
gacaaatgcg ttgaaaaccc aaacccaacg tgaatgaaa acaacggcgg ctgcgacgct 960
gacgccaaat gcacggaaga agacagcggc tcgaatggca agaaaattac ctgcgaatgc 1020
acgaaaccag acagctaccc actcttcgac ggcattttct tagt 1065

<210> 10
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:N-terminal
hexahistidine, His tag, affinity tag

<400> 10
His His His His His His
1 5